



Science & Technology
Facilities Council



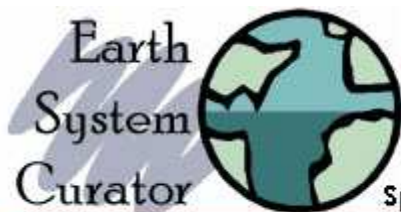
"finding and understanding simulations of past, present and future
climate"

Model Metadata for CMIP5

Presented by

B. Lawrence (BADC)

On behalf of the Metafor and Earth System Curator teams



Spanning the Gap Between Models and Datasets



Outline

➤ Context

- The METAFOR Project
- ESG Curator

➤ Documenting Models and Simulations

- The METAFOR Common Information Model (CIM)
- Controlled Vocabulary

➤ The “Questionnaire” for CMIP5

- Model Descriptions
 - Component Structure
 - Component Attributes
- Experiments Simulations

➤ Using the model metadata

➤ Summary



Discrimination and Documentation

“A problem for CMIP5”

CMIP5 will involve more than twenty different models running at least twenty institutions running dozens of simulations corresponding to dozens of numerical experiments.

The volume of data will be petabytes.
Even the core data will be of order $\frac{3}{4}$ of a petabyte.

What should one download?

What are the characteristics of the model that produced it?

How was a particular simulation constructed?

Rowan Sutton: “There are no end-users of climate predictions”

Gavin Schmidt: “the public ... will get used to dealing with climate model outputs ... however, an increased amount of hand-holding will be necessary”

But engagement requires effort and time!

... and it's not always the public who require “engagement” above and beyond “normal academic behaviour”, misunderstanding abound and even appear in the peer reviewed literature!



Metafor: Facts and Figures

Common Metadata for Climate Modelling Digital Repositories

<http://metaforclimate.eu>

**SEVENTH FRAMEWORK
PROGRAMME**



CAPACITIES

ures

INFRA-2007-1.2.1 - Scientific Digital
Repositories

Heritage: ENES:

- PRISM project (FP5 2001-2004)
- Followed by the PRISM Sustained Initiative (PSI)

11 partners

Total budget of 2.2M€

Started March 2008, duration 3
years

- NCAS Climate, NCAS CMS. Reading Uni, UK (Coordinator)
- NCAS BADC (STFC CEDA), UK
- CERFACS, France
- Models and Data, Max Planck Institute for Meteorology, Germany
- Institute Pierre-Simon Laplace, CNRS, France
- University of Manchester, UK
- Met Office, UK
- Administratia Nationala de Meteorologie, Romania
- Météo France, CNRM, France
- CLIMPACT, France
- CICS, Princeton University, USA



National Centre for
Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL



University of
Reading



M&D
Modells & Daten



MANCHESTER
1824



climipact
CLIMATE INFORMATION TECHNOLOGY



British Atmospheric
Data Centre

NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL

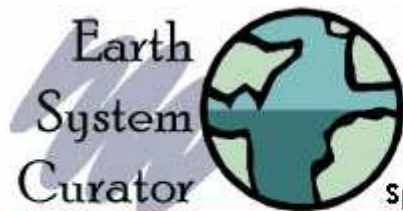
CONTEXT: EARTH SYSTEM CURATOR

GOALS The Curator project develops tools and portals that link simulations, models, components, datasets, and services in support of climate research and assessments. Curator workflows are driven by standardized metadata developed in collaboration with the Earth System Grid (ESG) and EU metafor projects.

PARTNERS NOAA GFDL, NCAR, Georgia Institute of Technology, Massachusetts Institute of Technology, plus external partners

PROJECTS

- **Metadata-enhanced science gateway** (lead NCAR)
Enhance the ESG portal with model metadata for improved understanding of datasets and their source models
- **Automatic code generation** (lead Georgia Tech)
Generate component couplers automatically using ESMF and the Eclipse Modeling Framework
 - **Climate modeling workflow** (lead GFDL)
Create a Flexible Run-Time Environment (FRE) that enables modelers to configure, execute, archive, reproduce, and modify climate model runs



Spanning the Gap Between Models and Datasets

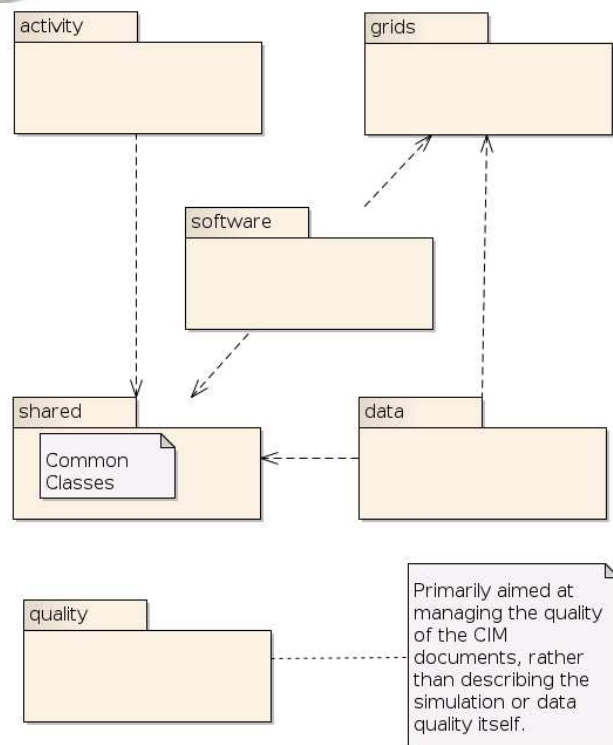


Slide Courtesy of Cecelia DeLuca , NCAR

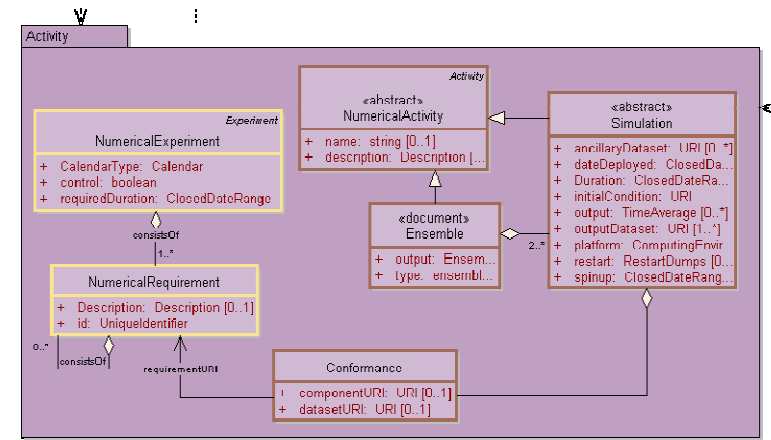
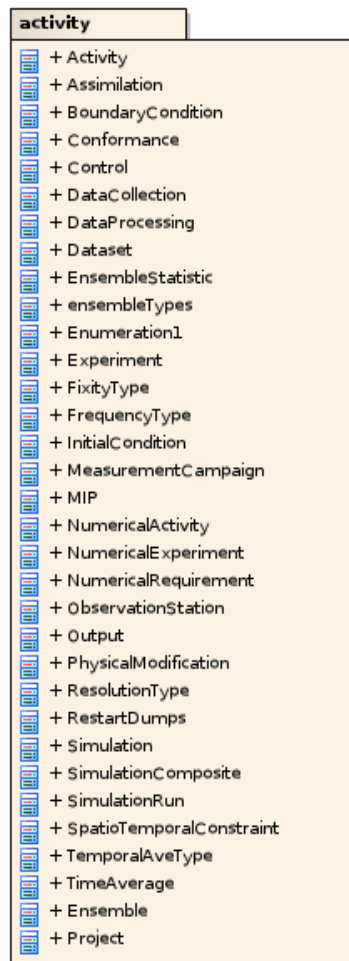
Curator is funded by the National Science Foundation and NASA



The Metafor Common Information Model



An activity uses software to produce data to be archived in a repository.

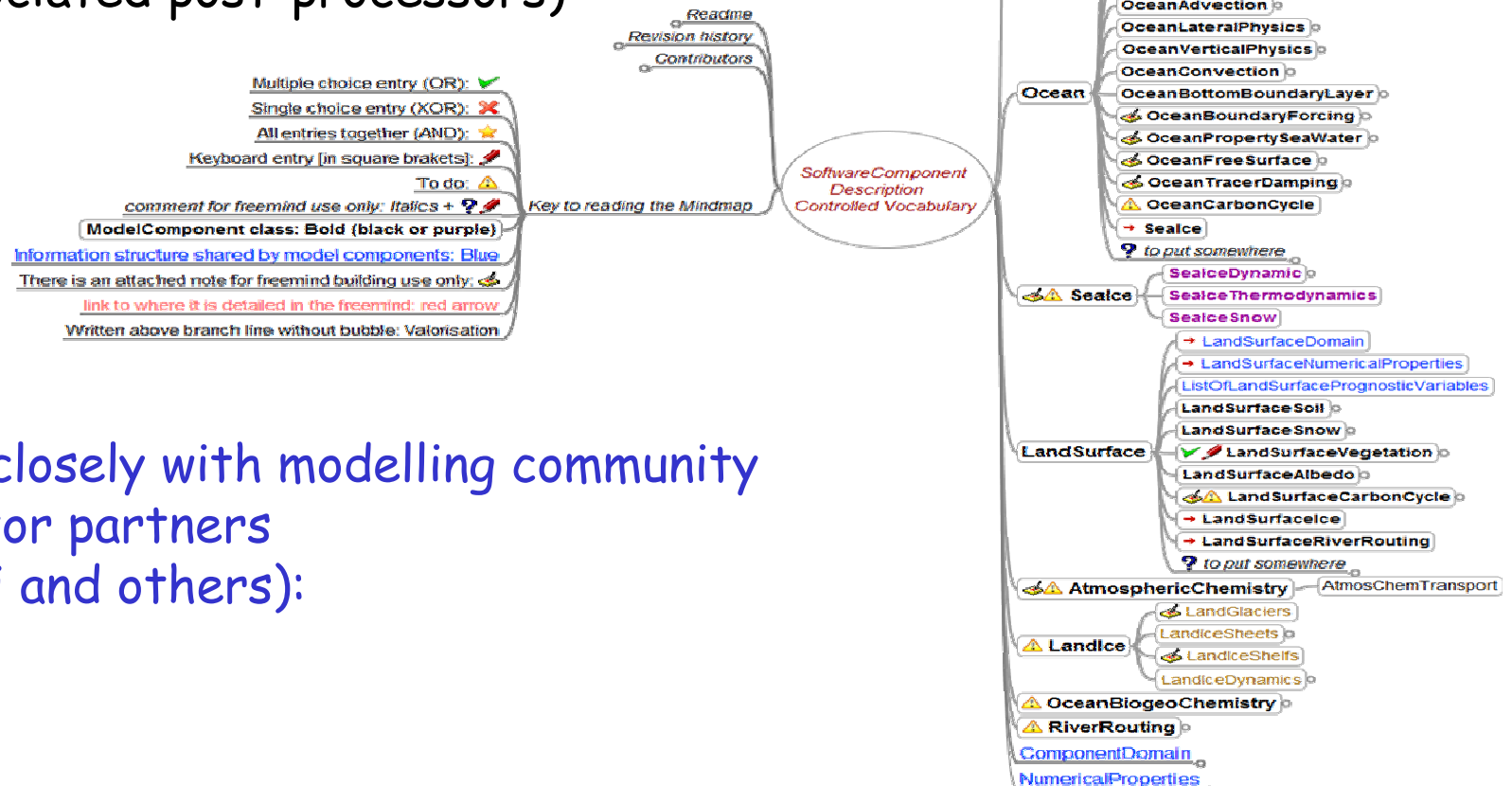


An simulation is a numerical activity which conforms to the numerical requirements of a numerical experiment.

Six packages with complex structure and various levels of interdependency.



Metafor developing controlled vocabularies to describe models (and associated post-processors)



Working closely with modelling community
and Curator partners
(via Balaji and others):



CMIP5 and Model Metadata

CMIP5:

- > 20 modelling centres
- > 50 numerical experiments
- > 86 simulations (total ensemble members) within experiments
- > 6500 years of simulation
- > Data to be available from “core-nodes” and “modelling-nodes” in a global federation.
- > Users need to find datasets, and discriminate between models, and between simulation characteristics.

CMIP5 Metadata Requirements:

- > Internal file metadata as required by CMOR.
- > Grid definitions using gridspec files.
- > Model Metadata as required by the “official” model metadata questionnaire ...
- > ...which is under active development. The current DRAFT version is available at <http://cmip5.metafor.ceda.ac.uk> (don't enter any real information in this version, it is being updated weekly and all data is deleted).



Questionnaire Screenshots, 1: Summary Page

The screenshot shows a web interface for a questionnaire. At the top is a navigation bar with links: Home:NCAS, HadBNL, Simulations, Files, References, Help, and About. The main heading is "Summary: UK National Centre for Atmospheric Science". Below this is an "Introduction" section with text explaining the purpose of the questionnaire and the entities it describes. Callouts point to specific parts of the page: "Centres can describe multiple models" points to the "Models associated with NCAS" section; "Centres can describe Input files and references." points to the "References" link in the navigation bar; "Centres can describe multiple platforms" points to the "Computing platforms associated with NCAS" section; and "Every simulation in the CMIP5 archive needs to be described." points to the "Simulations associated with NCAS" section.

Home:NCAS **HadBNL** **Simulations** **Files** **References** **Help** **About**

Summary: UK National Centre for Atmospheric Science

Introduction
Each CMIP5 modelling centre is running *Simulations* which run *Models* on *Platforms*.
The *Models* are made up of *Components*.
The *Simulations* conform to the *NumericalRequirements* of *Experiments* via what we call *Conformances* which consist of either specific code modifications or the use of specific boundary or initial condition *Files*.
The purpose of this questionnaire is to glean information about the entities denoted *thus* and/or their relationships.
We expect to see each centre enter at least one model, one platform, and then multiple simulations, each of which will involve entering descriptions of how they conform to the numerical requirements *Conformances*. It is not possible to start entering simulation information until at least one model and one platform have been created.

Models associated with NCAS

Model	Status
HadBNL	placeholder

[Add a new model](#) [Edit](#)

Note that it can take some time to create a new model from the CMIP5 template ... be patient!

The status column provides an indicator of how much of the model description has been completed.

Computing platforms associated with NCAS

Platform
Hector

[Add a new Platform](#) [edit](#)

Simulations associated with NCAS

Simulation
test

[Simulation Master Page](#)

Other

Number of dataobjects and conformances listed here

Questionnaire Screenshots

2: Model Description

Every model consists of components each of which are individually described

Home: NCAS | **HadBNL** | Simulations | Files | References | Help | About

Model Component HadBNL

All buttons and links in the tabs and this column navigate away from this page. Save your work first!

Available Models

- HadBNL
 - LandSurface
 - LandIce
 - Ocean
 - OceanBiogeoChemistry
 - Sealce
 - Atmosphere
 - AtmosChemAndAerosols

Component HadBNL

Please add details of any other relevant subcomponents of this component

Add Subcomponent

The button(s) in this box navigate to pages which further describe this **component**.

References

The buttons in this box navigate to pages for this **model**.

Coupling

View

Validate

Short Name: (type: , Implemented: ☒)

Long Name:

Contact Email:

Contact Name:

Component Attributes

In this section enter parameters and attributes associated with this component.

There are no required attributes at this level for this component. Your attributes:

g: (Delete ☐) Enter freetext or numeric value:

Enter an additional component attribute or parameter if you wish (you can enter multiple parameters by entering one at a time):

Name	Value
<input type="text"/>	<input type="text"/>

History

Year Released: Previous Version (name):

Please tell us how this component has been improved from that previous version.

Additional Information

Please provide additional information about this component as necessary.

Save

Arbitrary component attributes can be described

Model history is important



Questionnaire Screenshots

3: Component Description

Home: NCAS HadBNL Simulations Files References Help About

Model Component Radiation

All buttons and links in the tabs and this column navigate away from this page. Save your work first!

Available Models

- HadBNL
 - LandSurface
 - LandIce
 - Ocean
 - OceanBiogeoChemistry
 - Sealce
 - Atmosphere
 - DynamicalCore
 - Radiation
 - Longwave
 - Shortwave
 - TurbulenceConvectionCloud
 - OrographyAndWaves
 - AtmosChemAndAerosols

Component Radiation

Please add details of any other relevant subcomponents of this component

Add Subcomponent

The button(s) in this box navigate to pages which further describe this component.

References

Short Name: Radiation (type: Radiation) Implemented: ☒

Long Name:

Contact Email: joe@foo.bar

Contact Name: Joe Bloggs

Component Attributes

In this section enter parameters and attributes associated with this component.

	Value
AerosolTypes: Choose one or more of: Sea Salt	Urban Oceanic Sea Salt
GHG-Types: Choose one or more of: CO2	
RadiationTimeStep: Enter freetext or numeric: CO2	
Enter an additional component attribute or parameter (you wish (you can enter multiple parameters by entering one at a time):	
Name	Value

Additional Information

Please provide additional information about this component as necessary.

References (associated with this component)

Update this list using the modify reference button

Detailed component structure from controlled vocabularies

Detail for component attributes From controlled vocabularies

Still able to add your own attributes

Component level references



Questionnaire Screenshots

4: Experiment List

Home:NCAS HadBNL **Simulations** Files References Help About

Simulations

This page holds the master index of all the simulations run by NCAS, and relates them to the application. You are expected to use the add simulation button to add simulations, after which you can edit them. You can choose to add placeholders for all simulations now, or add them over time, and either way, come back and edit the simulation descriptions at a later date.

For each experiment, you can see listed the simulations from this centre conforming to the experiment.

(NB: The experiments, and their descriptions will be entered via the CIM instances)

name	simulations
CMIP5 Tier 1 Time-slice experiment 6.7a - AquaPlanet Control run	test Add New
CMIP5 Tier-2 future experiment 4.2L	Add New
CMIP5 Tier 2 Time-slice experiment 3.3E - AMIP Ensemble	Add New
CMIP5 Tier 1 Long-term experiment 3.5 - Last Glacial Maximum	Add New
Historical	Add New
CMIP5 Tier 1 near-term decadal experiment 1.5 - Alternative Initialization	Add New
CMIP5 Core near-term decadal experiment 1.1	Add New
CMIP5 Tier 2 Time-slice experiment 2.1E - Future "time-slice" experiment ensemble	Add New
CMIP5 Tier 1 Long-term experiment 3.6 - Last Millennium (850 - 1850)	Add New
CMIP5 Tier 2 experiment 7.1E - Natural-only ensemble	Add New
4.2 RepresentativeConcentrationPathway8.5	Add New
CMIP5 Tier 1 Time-slice experiment 6.4 - Hansen style response to aerosol	Add New
CMIP5 Tier 1 long-term experiment 3.2E - Historical ensemble	Add New
4.1 RepresentativeConcentrationPathway4.5	Add New
CMIP5 Tier 1 experiment 7.1 - Natural-only	Add New
CMIP5 Tier 1 Time-slice experiment 6.7c - 4K SST increase	Add New
5.1 EarthSystemModelControl	Add New
CMIP5 Tier 2 Time-slice experiment 6.8 - Cloud response to	Add New

All CMIP5 experiments fully described (see next slide)

Each simulation individually described

Each simulation associated with an experiment



Questionnaire Screenshots

5: Experiment Description

Experiment CMIP5_NumReq_6.7a

To explore the degree to which future climate states depend on the initial climate state (from 'A Summary of the CMIP5 Experiment Design' by Karl Taylor et al. 2008)

Duration 1850-01-01T00:00:00Z to 1856-01-01T00:00:00Z

Requirements

6.7a.ic	InitialCondition	A description of initial condition for the control run
6.7a.bc.wmg_anth.conc	BoundaryCondition	Imposed changing concentrations of historical anthropogenic well mixed gases, including CO2
6.7a.bc.wmg_volc.conc	BoundaryCondition	Imposed changing concentrations of historical volcanic well mixed gases, including CO2
6.7a.bc.sls.conc	BoundaryCondition	Imposed changing concentrations of historical short lived gas species
6.7a.bc.aer_nat.conc	BoundaryCondition	Imposed changing concentrations of historical natural aerosols
6.7a.bc.aer_anth.conc	BoundaryCondition	Imposed changing concentrations of historical anthropogenic aerosols
6.7a.bc.aer_pre_nat.conc	BoundaryCondition	Imposed changing concentrations of historical natural aerosol precursors
6.7a.bc.aer_pre_anth.conc	BoundaryCondition	Imposed changing concentrations of historical anthropogenic aerosol precursors
6.7a.bc.sls.em	BoundaryCondition	Imposed changing emissions of historical short lived gas species
6.7a.bc.aer_nat.em	BoundaryCondition	Imposed changing emissions of historical natural aerosols
6.7a.bc.aer_anth.em	BoundaryCondition	Imposed changing emissions of historical anthropogenic aerosols
6.7a.bc.solar	BoundaryCondition	Imposed historical changing solar forcing
6.7a.bc.aquaplanet_ssts	BoundaryCondition	Imposed zonally uniform SSTs on an aquaplanet

Placeholder for a list of case simulation numbers which have been documented for this experiment

Requirements for initial
and boundary conditions
preloaded in
questionnaire setup.

Questionnaire Screenshots

6: Simulation Description

Page layout
(tabs &
buttons)
will change
before release

Home: NCAS HadBNL Simulations Files References Update Help About

Update a simulation in CMIP5 NumReq 6.7a

(To explore the degree to which future climate states depend on the initial climate state (from 'A Summary of the CMIP5 Experiment Design' by Karl Taylor et al. 2008))

... buttons and links in the tabs
... and this column navigate away
from this page. Save your work
first

File Dependencies

Initial Conditions

Boundary Conditions

Experiment Criteria

Conformance

Ensemble Characteristics

Not an ensemble

Placeholder for validate
and view buttons

General Characteristics of the Simulation

In this box you fill out the general characteristics of the simulation (or ensemble).

Short Name: test

Long Name: adsf

Contact Email: adf@foo.bar

Contact Name: adsf

Simulation used: Model HadBNL

Ensemble size: 1

Description (notes)

Update

Other pages
describe initial &
boundary conditions
& conformance ...

... and how
those
correspond
to the
experiment
requirements

Need to
complete
Initial and
Boundary
conditions

Initial and Boundary Conditions

(You can modify these using the buttons on the left hand menu.)

You need to set an initial Condition

Boundary Condition: Coupling Atmosphere:SST (for test) uses 0 internal link(s) and 0 link(s) to files. Incomplete

(Each boundary condition must have at least one internal or external coupling.)

Conformance: How this simulation conforms to the numerical requirements of experiment CMIP5_NumReq_6.7a

(You can modify these using the button on the left hand menu.)



Summary

- CMIP5 Model metadata requirement in response to large international projects capturing details of what is needed.
- Considerable flexibility in what can be captured, but much is required too.
- The “questionnaire tool” to collect the metadata will change considerably in the near future, but will be available by late (NH) autumn.
- The archive centres intend to mandate the requirement for this metadata before making data available to community.
- ESG and Metafor/IS-ENES will deploy catalogues to exploit this metadata, linking data to information about the data allowing a wider community to accurately exploit the considerable investment in CMIP5 design and execution.